Transcriptional and Chromatin Dynamics of Muscle Regeneration After Severe Trauma

Carlos A. Aguilar[1,*], Ramona Pop[2], Anna Shcherbina[1], Alain Watts[1], Ronald W. Matheny[3], Davide Cacchiarelli[2], Christopher T. Carrigan[3], Casey A. Gifford[2], Melissa A. Kottke[3], Woojin M. Han[4,5,6], Eun Jung Shin[5,6], Young C. Jang[5,6], Jackson Lee[1], Maria L. Urso[3,†], Alexander Meissner[2]

[1]Massachusetts Institute of Technology - Lincoln Laboratory, Lexington, MA 02127. [2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, Harvard Stem Cell Institute, Cambridge, MA 02138, Dept. of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138. [3]United States Army Institute of Environmental Medicine - Military Performance Division, Natick, MA 01760. [4]Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332. [5]School of Applied Physiology, Georgia Institute of Technology, Atlanta, GA 30332. [6]Parker H. Petit Institute of Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332.[†]Present address: Smith and Nephew, Biotherapeutics, Ft. Worth, TX, 76132.

*To whom correspondence should be addressed: carlos.aguilar@ll.mit.edu

Abstract: Following injury, adult skeletal muscle undergoes a well-coordinated sequence of molecular and physiological events to promote repair and regeneration. However, a thorough understanding of the *in-vivo* epigenomic and transcriptional mechanisms that drive and control these reparative events is lacking. To address this, we monitored the *in-vivo* dynamics of three histone modifications and coding and noncoding RNA expression throughout the regenerative process in a mouse model of traumatic muscle injury. We first illustrate how the transcriptional landscape of coding and non-coding RNAs (both long non-coding RNAs and microRNAs) in tissues and sorted satellite cells is modified and regulated during various stages after trauma. Next, we use chromatin immuno-precipitation followed by sequencing (ChIP-Seq) to evaluate the chromatin state of *cis*-regulatory elements (promoters and enhancers) and view how these elements evolve and influence various muscle repair and regeneration transcriptional programs through different stages of healing. These results provide an unbiased, comprehensive view of the central factors that regulate *in-vivo* muscle regeneration and underscore the multiple levels through which both transcriptional and epigenetic patterns are regulated to enact appropriate repair and regeneration.

Introduction

Adult skeletal muscle is a post-mitotic organ that coordinates movement and constantly grows and adapts by remodeling its structure and metabolism. After insult or injury, adult skeletal muscle enables repair and regeneration of existing fibers and formation of new fibers through a population of stem cells that reside underneath the basal lamina called satellite cells[1] (SCs). The SCs reside in a specialized *niche*[2,3] and change their quiescent complexion after injury, whereby SCs activate, proliferate, differentiate into myoblasts and fuse to form myofibers. Each step is affected by various environmental signals and communication with infiltrating[4,5] and resident cells.[6] It is well established that multiple repair and regeneration sub-processes accomplished by SCs (and other cell types) after muscle injury are orchestrated by distinct transcriptional networks,[7] encompassing epigenetic,[8,9] transcriptional[10] and post-transcriptional events. However, the integrative dynamics of transcriptional networks and regulatory epigenetic switches at genome-wide levels have not been characterized *in-vivo* and as such, our understanding of the molecular processes and transcription factors (TFs) involved in muscle regeneration have been limited. Profiling the expression and chromatin state of *cis*-regulatory elements[11] after severe muscle trauma provides a powerful method to understand the molecular determinants of healing progression and can provide crucial insights for development of therapeutic modalities for effective repair and regeneration of severe muscle trauma, myopathies such as Duchenne muscular dystrophy and aging.

The *cis*-regulatory networks that orchestrate *in-vivo* muscle repair and regeneration after traumatic injury have only been partly characterized[12,13] and herein, the *in-vivo* evolution of coding and noncoding expression and three different chromatin modifications (H3K4me3, H3K4me1 and H3K27ac) were profiled across nine time points (t = 3 hrs to 672 hrs) from an injured and uninjured contralateral tibialis anterior muscle. The generated genomic maps were then contrasted against MyoD and MyoG genomic-binding data[14,15] to determine shared and distinguishing signatures at *cis*-regulatory elements during different stages after injury. The dynamic levels of numerous coding and noncoding transcripts, chromatin state transitions, and differential binding at TF motifs were integrated and assessed to construct a comprehensive view of the key transcriptional and chromatin factors that influence and modulate *in-vivo* muscle repair and regeneration dynamics.

Results

Severe muscle trauma induces extensive transcriptional and post-transcriptional regulation in both coding and noncoding transcripts

To better understand the evolution of the healing process after administration of trauma, expression profiling by RNA sequencing[16] (RNA-seq) and small RNA-Seq (miRNAs) was performed for both the injured and contralateral tissues for multiple time points (Fig. 1). Hierarchical clustering of the RNA-Seq data through time revealed clusters up- and downregulated at different time periods that were associated with different stages of muscle repair and regeneration (Fig. 1b). For example, chemokine ligands 2 and 7 (CCL2 and CCL7), which are important for the recruitment of various immune cells to the injured muscle, peaked in expression in the early period (at 24hrs after injury). Annexins 1 and 2 (Anxa1 and Anxa2), cellular membrane binding proteins that are important for promoting migration of satellite cells, demonstrated a different expression profile and peaked in the middle period (between 48hrs and 72hrs) and subsided back to control levels approximately two weeks after injury. Laminin

subunits beta-1 and alpha-4 (Lamb1 and Lama4), which are cellular adhesion molecules present in the basement membrane, exhibited a temporal expression profile that peaked in the late period (at 336hrs after injury). Previously, we performed pathway analysis of the time-clustered RNA-Seq data[16] and showed an initial burst of pro-inflammatory and immune-response transcripts in the early period, followed by activation, proliferation and differentiation of myogenic precursors and extra-cellular matrix (ECM) remodeling in the middle and late time periods, which was consistent with previous studies of muscle tissue injury and in line with a productive healing process.[17,18]

Positive and negative regulation of muscle regeneration has also recently been shown to occur post-transcriptionally, where microRNAs (miRNAs) bind to the 3′ untranslated region (UTR) of mRNAs and inhibit translation.[19,20] Small RNA-Seq was performed at multiple time points from each stage (early: 3hrs & 10hrs; middle: 72hrs, 168hrs; late: 336hrs, 672hrs) and 841 miRNAs were detected for at least one time point. Of the 841 expressed miRNAs, 143 showed dynamic behavior (See Methods) and analysis of the dynamic miRNAs reinforced many of the results observed from the RNA-Seq datasets, where early upregulated miRNAs were associated with inflammation and immune system programs and middle and late stage upregulated microRNAs were associated with muscle repair and regeneration (Supp. Fig. 1). For example, in the early period, the inflammatory miR-223 was upregulated as to where its expression subsided in the middle and late periods. In the middle period, a new subset of miRNAs were upregulated such as miR126b (associated with angiogenic and chemokine signaling[21]), the miR-29 family (regulator of fibrosis and collagen expression), miR-21 (regulator of PI3k/Akt signaling), miR-31 and miR-206 (differentiating myoblasts), which was consistent with pathways observed promoting myogenic differentiation[16,22,23] during the middle period.

To gain insights into the consequences of increases in expression of these dynamic miRNAs, we assessed pairwise correlations in expression patterns with their 3' UTR mRNA targets (see Methods). Uniquely, we observed multiple miRNA and mRNA targets that were reciprocally regulated during the time-course and hierarchical clustering of the correlations revealed 3 clusters (Fig. 2a). The first cluster contained positively correlated miRNA-mRNA targets that were significantly upregulated in the early period and could be ascribed to functional classes related to immune regulation and DNA damage cellular response such as miR-709, miR-6538, miR-378 family with genes Atf3, Stat5b, Fas and Btg2. The second cluster of positively correlated miRNA-mRNA targets were upregulated in the middle and late periods and were associated with myogenic growth and development. For example, we detect upregulation of miR-206, miR-205, miR-203, miR-1a, miR182, and miR-31 along with several target genes that were previously identified as critical regulators of satellite cell differentiation[24] (Igf2, Igfbp5) in the middle period, further suggesting the onset of differentiation during this time. The third cluster possessed the largest number of positively correlated miRNA-mRNA targets and these transcripts were also upregulated in the middle period. This cluster was linked to TGF-β signaling and ECM and cytoskeletal remodeling and contained miR-18a, miR-21, miR-335, miR34c and miR-142 along with target genes Tgfbi, Anxa1, Sdcbp, Cav1 and Timp2.

Long non-coding RNAs (lncRNAs) are a new class of transcripts that have recently been discovered and play diverse regulatory roles in muscle differentiation and disease.[25,26,27] Previously identified lncRNAs were intersected with the generated RNA-Seq data and 2444 lncRNAs were detected, with 124 lncRNAs exhibiting dynamic differential expression ($\log_2$ Fold Change > 2 and adjusted p-value (FDR) < 0.1) for at least one time point. Figure 2b shows the

differentially expressed lncRNAs clustered into 3 groups and similar to the mRNA and miRNA datasets, we view upregulated lncRNAs involved in inflammatory and immune-related processes (Csf1r, Il1bos, Ly9, Treml1, CD68, Ncf1) in the early time points followed by increases in expression of lncRNAs that were previously shown to be critical for myogenic differentiation[26,28] (H19 and linc-MD1) in the middle and late periods which was also coincident with the peak times observed for the onset of myoblast differentiation found from RNA-Seq and miRNA-Seq data. Uniquely, in the early period we also detected upregulated lncRNAs that were found in other injured muscle tissues[29] (myocardial infarction-associated transcript – MIRT1 & 2), suggesting a common regulatory scheme.

Temporal transcriptional signatures of sorted satellite cells reflects a dynamically changing tissue composition

Satellite cells from injured tissues were isolated using flow cytometry (Sca-1$^-$, CD45$^-$, Mac-1$^-$, Ter-119$^-$, β1-integrin$^+$, Cxcr4$^+$, Fig. 3a) at time points associated with activation, proliferation and differentiation (3, 24, 48, 72 and 168 hrs) and expression profiling by RNA sequencing (RNA-seq) and small RNA-Seq (miRNAs) was performed. 17,636 mRNAs were detected (TPM > 1) and 9,953 genes were differentially expressed (fold change > 2) at one or more time points (Fig. 3b). Of the time points sampled, 168 hrs possessed the highest number of genes that were differentially expressed (584) compared with 434 genes for 3 hrs, 108 genes for 24 hrs, 405 genes for 48hrs, and 241 genes for 72 hrs. These datasets were compared to previously published microarray datasets[12,30,31] of isolated satellite cells and excellent agreement was viewed. Hierarchical clustering of the datasets through time revealed 5 clusters, and gene ontology (GO) analysis of the differentially expressed genes within each cluster illustrated significant pathway perturbations (FDR < 0.05) in all five clusters. (Fig. 3b). A variety of upregulated genes that remained overexpressed over the course of the times sampled were in third cluster and related to ATP synthesis, ion channel activity and regulation of glucose and insulin. Multiple genes associated with the fifth cluster were common markers of quiescent satellite cells (Pax7, Foxo3) and displayed a relatively high level of expression from 3-48hrs that decreased from 72-168 hrs (Fig. 3c). A different set of genes was associated with the second and fourth clusters (Fig. 3c) that were low in expression from 3-48hrs and upregulated from 72-168 hrs. These two clusters contained multiple genes associated with myogenic differentiation (MyoD1, MyoG, Des, Ash2l, Hes6), cell cycle regulation (Ccna2, Ccnb1, Dnmt1, Birc5), cytoskeletal proteins (Cdh15, Tmem8c/Myomaker) and mitochondrial metabolism.

miRNA-Seq was performed on the FACS-sorted satellite cells and 341 miRNAs were detected for at least one time point (Fig. 3d). 107 detected miRNAs were differentially expressed (Supp. Fig. 1) and many of the detected miRNAs mirrored expression patterns observed from whole tissues. For example, in the early period, miR-22 was downregulated in both satellite cells and at the tissue level. miR-22 has recently been shown[32] to inhibit HDAC4 expression, which is a well-known negative regulator of myogenesis. During this time period, the miR-181 family (modulators of PI3k signaling and metabolic adaption needed for subsequent proliferation) and miR-191 (promotes cell migration via induction of transforming growth factor beta (TGFβ) signaling) also showed decreases in expression, which is consistent with transcriptional changes in metabolism, TGFβ-signaling and PI3k signaling observed in the tissue at that time. In the middle time periods (≥72 hrs), changes in expression of let-7c (SMAD and TGFβ signaling inhibitor), miR-16 (regulator of proliferation[33] and cell-cycle genes), miR-148a (modulates fate commitment), miR-182 (repressor of negative regulators of cell-cycle genes) and several other

members of the let-7 family (associated with cell cycling and self-renewal) were detected compared to low expression in the earlier time points (3-48 hrs). miR-486a and miR-486b were also upregulated at these time points, and were previously shown to potentiate proliferation by targeting p85α, IGF1R, and IGF1, which is consistent with a decrease in expression of those genes at the same time points (72 and 168 hrs). Multiple other miRNAs exhibited similar expression profiles (low expression from 3-48hrs and high expression from >72 hrs) such as 21a, 125a, 127, 199a, 206, 411, 541 and were consistent with previous findings[34] for isolated satellite cells.

Acute muscle trauma stimulates genome-wide chromatin remodeling

To further probe into the regulation of the coding and noncoding transcriptional programs, chromatin immunoprecipitation followed by DNA sequencing (ChIP-Seq) was used to globally map the chromatin state of various *cis*-regulatory elements at nine time points that were lumped into three key stages after a traumatic muscle injury (early: 3-24 hours after injury, middle: 48-168 hours after injury, late: 336-672 hours after injury). Validated antibodies for histone H3 lysine 4 trimethylation (H3K4me3), a modification associated with promoters, H3K4me1 and H3K27ac, associated with poised and active enhancer regions[35], respectively, were used to enrich chromatin. Each tissue from each time point was immuno-precipitated using the three antibodies, and ChIP-Seq maps from the same antibody and time point were merged, resulting in 54 chromatin-state maps covering >1.5 billion reads.

In the early time period, 93,149 sites were enriched for H3K4me3, 104,890 sites for the middle period followed by 141,948 sites for the late period (Fig. 4). Of these, 1,606 were differentially enriched in the early time points relative to the controls, 6,733 in the middle time points, and 6,121 in the late time points. To integrate sites that gained or lost the H3K4me3 modification during the time course with transcriptional activity (FPKM>1), RNA sequencing (RNA-Seq) results were merged with H3K4me3 enrichments. Approximately 58% of sites were found to associate with transcriptional activity (FPKM>1) and 4,241 sites exhibited dynamics (acquisition or loss of the histone modification) through at least one stage (early-middle, middle-late, early-late).

In contrast to H3K4me3, which has previously been shown to be largely static during chromatin remodeling events,[36] H3K4me1 and H3K27ac, which demarcate enhancer elements, were observed to be highly dynamic.[37] H3K4me1 resides on both poised and active enhancers, as to where H3K27ac marks active enhancers.[38] In the early time period, 93,938 sites enriched for H3K4me1 were identified, 106,353 sites for the middle period, and 28,663 sites for the late period. The active enhancer mark H3K27ac showed similar behavior with 32,285 sites enriched in the early time period, 43,780 sites in the middle period and 13,748 sites for the late period. The highest number of sites that acquired H3K27Ac was found in the 72hr period (28,178 sites) compared to 11,535 for the 48hr stage and 4,067 for the 168hr stage (Fig. 4). In contrast to H3K4me3 enrichments, which were primarily found at or near transcriptional start sites, the enhancer marks H3K4me1 and H3K27ac were more broadly distributed across inter- and intragenic loci (Fig. 4). Since H3K4me1 demarcates both poised and active enhancers and H3K27ac marks only active enhancers, the total number of enhancer elements (identified by $H3K4me1^{+}$, $H3K27ac^{+}$ and $H3K4me3^{-}$) was determined and the ratio of the enriched enhancers was quantified as active ($H3K4me1^{+}$ & $H3K27ac^{+}$) or poised ($H3K4me1^{+}$ & $H3K27ac^{-}$). A wide spectrum of binding patterns was observed for the two enhancer categories[39] across the nine time points with the majority of differential active enhancers occurring at intragenic and intergenic

loci for all time periods. The highest number of differential active enhancers (injured^{+} & control^{-}) was found for the middle period, with the largest number of active enhancers occurring in the 72hr period.

Integrative analysis of enriched enhancers highlights transient recruitment of different types of immune cells after muscle trauma

In the early period, numerous transcripts associated with inflammation, invading immune cells[4], cytokine signaling, and apoptosis were detected. GO term analysis of the enriched enhancer peaks for the early period for the injured samples compared to uninjured controls demonstrated overrepresented terms similar to the transcriptional groups (immune response, chemokine and cytokine signaling, inflammation, and metabolism of lipids and lipoproteins, Fig. 5a). The result of these chromatin and transcriptional enrichments in the early period is consistent with upregulation of cytokines and chemokines (IL-6 and TNFα),[16] and increases in expression of pleiotropic transcription factors[36] such as components of the AP-1 complex (c-Fos, Atf3 and JunB), STATs, NF-κβ, EGR and PU.1 (Fig. 5b and Supp. Fig. 2), which stimulate a permissive chromatin state and have previously been shown to induce satellite cell activation genes such as Myod1, Myf6 and c-Myc.[40]

In the middle time period (48-168hrs after injury), multiple genes associated with TGF-β signaling were differentially expressed, such as several SMADs (Supp. Fig. 3), which have been shown to interact with chromatin remodeling complexes such as histone acetyltransferases p300 and CBP (CREB-binding protein) to induce H3K27 acetylation.[41] Consistent with this view, we observe the highest number of sites that acquired H3K27ac was in the middle time period (Fig. 4). GO annotation of the differential enhancer peaks for the middle period revealed multiple enrichments for anti-inflammatory cytokines, G-protein coupled receptors and the Rho family of small GTPases (Fig. 5 and Supp. Fig. 4). This result is in agreement with the peak in expression of SRF target genes[42] and anti-inflammatory cytokines interleukin-4, 10 and 13 (IL-4, IL-10, IL-13), which induce M2 macrophage polarization and are essential components for resolution of inflammation and tissue repair[5] (Fig. 5). Increases in expression of these factors was also mirrored by upregulation of the macrophage-derived matrix metalloproteinase 12 (Mmp12), which cleaves and inactivates CXC chemokines (Cxcl1, -2, -3, -5, and -8) and monocyte chemotactic proteins (Ccl2, -7, -8, and -13), inhibiting leukocyte flux to the injured site and abrogating the amount of pro-inflammatory molecules present in the injured tissue (Supp. Fig. 5).

Chromatin state transitions associated with activation of ECM repair & myogenic regeneration

A primary determinant of successful muscle regeneration after injury is the remodeling of extracellular matrix[43] (ECM), which provides a scaffold for proliferating satellite cells to migrate, differentiate and fuse in the correct orientation. In the middle period, increases in expression and multiple enhancer enrichments were viewed for genes associated with ECM remodeling, such as components of the basement membrane (myoferlin, laminin, collagen VI genes, annexins, Fig. 6 and Supp. Fig. 5), glycoproteins and matrix metalloproteinases (MMPs). Additionally, during the middle period, chromatin remodeling and upregulation of genes associated with muscle development and architecture were also detected (Fig. 6).

In order to understand this critical step further, additional analysis of the differential active enhancer sites during the middle period was conducted by evaluating the number of enriched enhancers that overlapped with MyoD and MyoG ChIP-Seq data.[14,15,44] The largest increase in

MyoD binding sites at active enhancers was found at 72hrs (Supp. Fig. 9), which is consistent with the largest increase in expression (Fig. 6e) and known association between MyoD and p300/CBP on E-box motifs of target genes during myogenic differentiation.[45] The greatest increase in MyoG binding sites at active enhancers was viewed at 336 hrs, which agrees well with observation of late myogenic differentiation programs at that time period. Compiling the enriched enhancer sites coincident for MyoD or MyoG binding and performing GREAT analysis revealed over-represented terms nominally associated with myogenic differentiation such as different types of growth factor signaling (insulin growth factor – IGF, fibroblast growth factor – FGF, hedgehog), p38 mitogen-activated protein kinase (p38-MAPK), and phosphatidylinositol 3-kinase (PI3K) signaling (Supp. Fig. 6 and discussed further below).

Chromatin regulation of PI3K pathway reflects transition from myoblast proliferation and differentiation

IGF signaling plays a crucial role in muscle repair and regeneration after injury, most notably during the transition between myoblast proliferation and differentiation where IGF2 activates the Akt pathway and targets bound by MyoD. In line with this observation, an increase in expression of IGF-1 was detected in the early period,[46] which subsided in the middle period when several IGFBPs and IGF2 were observed to be upregulated along with miR-483 (which is embedded in the second intron of IGF2 and is an SRF target) and members of the Akt pathway (Fig. 7a).[47] One component of the IGF pathway, the PI3K component, is essential for skeletal muscle regeneration and composed of three primary classes (Class I, II, III) that are structurally and functionally distinct. The three Class IA PI3K p110 catalytic subunits (α, β, and δ) are expressed in skeletal muscle[48] and of these subunits, p110α and p110β have been shown to positively influence myoblast proliferation and differentiation.[49,50] In the early and middle periods following injury, we observed differential enhancer binding and increases in expression of PI3K catalytic subunits p110α (*Pik3ca*), p110β (*Pik3cb*), and p110δ (*Pik3cd*) as well as the PI3K p85α (*Pik3r1*) regulatory subunit for the injured muscle as compared to the contralateral control (Fig. 6a). Increases in expression and active enhancers were also viewed for Class IB p110γ (*Pik3cg*) and its regulatory subunit p101 (*Pik3r5*) during the healing process, which are expressed abundantly in immune cells. TF-binding analysis of the enriched enhancer regions for the regulatory subunit p85α showed motifs for signal mediators of IL-4 and the anti-inflammatory response, which is consistent with activation of pathways during this time period (Fig. 7b).

The temporal increases in PI3K expression (and other positive myogenic regulatory pathways) between 24-72hrs post-injury were also concomitant with increases in expression of negative feedback regulators such as MG53 (disruptor of upstream IGF signaling via ubiquitin ligases)[51], XBP1[52] and USF1 (which compete with or co-occupy binding sites for MyoD and Mef2), and Id1, Id2 and Snai1[53] (which reduce the binding affinity of MyoD and Mef2). The negative regulators began to decrease in expression starting at 72hrs (Supp. Fig. 7), further reinforcing the observation of induction of myogenic differentiation at this time.

Discussion

Muscle recovery after traumatic injuries such as an occupational crush or blast suffered during military combat canonically induces a predisposition for additional injuries and chronic pain owing to incomplete regeneration of the tissue. Many of the various muscle repair and regeneration sub-processes triggered after trauma have been studied, yet many questions at the

molecular level remain such as how the *in-vivo* epigenomic landscape evolves. Accordingly, our understanding of how different factors translate to influence the chromatin architecture that regulates expression at lineage-specific genes has been limited. Herein, we used integrative genomic mapping technologies to profile coding and noncoding expression and the *in-vivo* chromatin state of various *cis*-regulatory elements and found successive waves of transcriptional and chromatin changes during the course of healing. To obtain broad, unbiased views of the diverse repair and regeneration pathways utilized by different cell types after trauma, we performed profiling of both whole, unfractionated muscle tissue and FACS-sorted satellite cells.[54] Whole tissue profiling enabled capture of signals emanating from multiple cell types (such as macrophages, fibro-adipogenic progenitors, fibroblasts, and endothelial cells), as well as the capacity to profile multiple chromatin modifications from a single tissue. Whole tissue profiling also permitted analysis of multiple types of histone modifications associated with enhancer elements without pooling tissues from multiple animals. The contralateral / uninjured tissue was used as the control for this study, however recently[55] it was shown that systemic signals induced from acute trauma stimulate satellite cells within the contralateral muscle to undergo a transition to an alert state that primes their differentiation potential. Thus, while we normalized changes in expression and chromatin remodeling to the contralateral tissue, the normalization may not be reflective of an unperturbed muscle.

Pathway analysis of the time-clustered whole-tissue RNA-Seq data[16] revealed waves of transcription associated with pro-inflammatory and immune-responses in the early period, which overlapped with transcriptional signatures associated with activation and proliferation of myogenic precursors. As these signatures subsided in the middle and late time periods, another wave of transcripts increased in expression that were associated with myogenic differentiation and extra-cellular matrix (ECM) remodeling. Collectively, these results are in line with a productive healing process.

Many noncoding RNAs (ncRNAs) play critical roles in regulating skeletal muscle gene expression programs and disruption of the muscle integrity by acute trauma produced coding and noncoding (miRNA and lncRNA) dynamics very similar to those observed in other muscle myophathies,[23,25,56] suggesting the modulation of common molecular pathways. A unique feature of ncRNAs is their ability to hybridize to other mRNAs and perturb their ability to be translated. We used miRanda-mirSVR[57] to study the interaction of dynamic miRNA-mRNA pairs and if their expression patterns change concomitantly with time. We identified 200 miRNA-mRNA mutually dynamic relationships and clustered the pairs into three categories of immune regulation, TGF-β signaling and ECM and cytoskeletal remodeling followed by myogenic differentiation. These three clusters highlight a temporal regulatory program where infiltrating immune cells release signaling molecules that trigger satellite cell activation followed by a transition whereby the activated progenitors proliferate and are repressed from differentiating by TGF-β signaling and changes to the surrounding matrix. These repressive cues were further extended via other regulatory mechanisms (see below) and began to subside at 72hrs when a new class of ncRNAs (miR-206, miR-205, miR-203, miR-1a, miR182, miR-31, H19, linc-MD1) increased in expression.

Long noncoding RNAs (lncRNAs) play diverse roles in muscle regeneration and among these roles are sequestration of miRNAs that block myogenic differentiation[26] (such as linc-MD1 with miR133), and regulation of cell-cycle exit genes[27] that mediate myogenic differentiation. Similar to the mRNA and miRNA-Seq results, three clusters of lncRNAs were detected with

inflammatory and immune-related lncRNAs upregulated in the early time points. In the middle and late periods, the inflammatory and immune-related lncRNAs decreased in expression while a new set of lncRNAs that were previously shown to be critical for myogenic differentiation (H19 and linc-MD1) increased in expression. Interestingly, H19 (which was upregulated in the middle and late periods) and another lncRNA called lncMyoD (which was upregulated in the early and middle periods) bind mRNA IGF2-binding proteins[27,28], suggesting induction and stabilization of this pathway may be regulated temporally via different combinations of lncRNAs. Additional regulation of the IGF pathway was viewed through induction of other negative modulators[51,52] (MG53, XBP1), further indicating regulation of this pathway extends across multiple levels and temporally. Since muscle repair and regeneration utilizes many feed-forward and feedback loops, tightly regulated expression patterns across multiple levels may facilitate a precise way to prevent extrinsic signal propagation from adjacent tissues that are also regenerating or responding to the injury, reinforce or tune noisy expression patterns, and perhaps facilitate a metabolically efficient mechanism to quickly respond after injury.

Expression profiling of sorted satellite cells for both coding and noncoding RNAs reinforced transcriptional dynamics observed from whole tissues, whereby in the early period, immune-response and chemokine transcripts were upregulated subsequently followed by proliferation and differentiation of satellite cells during the middle time periods. Several TFs associated with quiescent cells (Pax3, Pax7, Foxo3, Spry1) and immune response (c-Jun, c-Fos, Atf3, Egr1) exhibited downregulation at the same time that genes associated with cell cycle progression (Ccna2, Ccnb1, Dnmt1, Birc5), differentiation and fusion (MyoG, Hes6, Tmem8c, Cdh15) were upregulated (>72hrs). Uniquely, at the same time period (72-168hrs) we also observe a switch from high-to-low expression of lysine demethylase 6A (KDM6a or UTX), which regulates removal of the repressive chromatin mark (H3K27me3),[58] to low-to-high expression of a component of a histone methyltransferase (HMT) complex (Ash2l). Ash2l is part of the Trithorax complex that specifically methylates the fourth lysine residue of histone H3 (H3K4) and is targeted by Pax7 to activate Myf5 target genes.[59] Other chromatin regulators such as PRDM12 (Combining these dynamics suggests in the early time period, upregulated TFs and chromatin remodelers aid to establish a more accessible chromatin state at myogenic loci, which is consistent with the increase in enrichments of H3K4me3, H3K4me1 and H3K27ac. Moving into the middle period, as stress-responsive and other immune-activated transcripts decrease in expression, another set of TFs and chromatin remodelers is then upregulated and titrated via miRNAs (see below) to enact appropriate satellite cell proliferation and differentiation.

miRNA-Seq on the FACS-sorted satellite cells demonstrated that the programs of activation, proliferation and subsequent differentiation are tightly controlled. For the time points linked with activation and mobilization (3-48 hrs), several microRNAs were downregulated such as miR-22, which is activated by the transcription factors AP-1 and NF-κB.[32] Since miR-22 inhibits expression of HDAC4 and targets c-myc, which also suppresses MyoD-initiated myogenic differentiation and promotes proliferation, a regulatory program can be inferred whereby stimuli from infiltrating immune cells induce activation of TFs that promote the progenitors to activate, mobilize and begin to proliferate but are repressed from differentiating. Other post-transcriptional factors such as miR-181, miR-191, miR-222 appear to reinforce this program in the early time periods and as inflammation and immune-stimulated signatures drop in expression, another program begins around 72hrs via increases of expression of regulators of

cell-cycle genes (miR-16, miR-182, miR-486) and myogenic differentiation (let-7 family, miR-206).

Repair and regeneration of skeletal muscle has canonically been described through the sequential activation of myogenic transcription factors such as MyoD, MyoG and the Mef2 family. The efficacy of these factors to initiate and maintain gene expression programs depends critically on the chromatin state[60,61] of their targets as well as their interactions with each other[62,63] and other TFs. Quantifying the genome-wide changes in histone modifications revealed a robust induction of chromatin remodeling occurs at a genome-wide level after muscle trauma and *cis*-regulatory elements such as promoters were largely invariant when compared to regions demarcated by enhancer marks (H3K4me1, H3K27ac). Immediately after the injury, the enriched enhancer regions were associated with canonical mediators of early stress, immunity and growth factor responses, which was consistent with the upregulation of cytokines and TFs such as AP-1 (Fos/Jun), SRF, NF-κβ, EGR, and STATs. The detection of these transcripts was expected since an appropriate inflammatory response is essential for functional recovery after muscle injury. A significant fraction of the enriched chromatin sites correlated with transcriptional dynamics and was most likely attributable to infiltrating monocytes that secrete and respond to cytokines and chemokines. These early immune-related changes were serially followed by changes in expression and chromatin remodeling of genes associated with anti-inflammatory macrophages[64] and activation and proliferation of satellite cells. During this period, upregulation of SMADs and TGF-β signaling was observed along with increases in expression of TFs such as Tead4,[65] Runx1[66] and MyoD[67]. SMADs have been shown to interact with chromatin remodeling complexes such as histone acetyltransferases p300 and CBP (CREB-binding protein) to induce H3K27 acetylation[68], which was consistent with the observation that the highest number of sites that acquired H3K27ac was in the middle time period. Additionally, MyoD has been shown to associate with p300 and p300/CBP on E-box motifs of target genes[69] and interactions between MEF2 and p300 enhance myogenic differentiation[70], suggesting a cooperative interaction between these different TFs and chromatin remodeling complexes during this phase of the injury response. Integrating these results reveals a transient shift in the balance of pro- and anti-inflammatory cytokine programs (high expression of IL-6 and low expression of TGF-β in the early period to low IL-6 expression and high expression of TGF-β in the middle period), which has previously been shown to promote the recruitment and differentiation of different types of cells that potentiate muscle repair[5] (neutrophils, M1 macrophages, T_H17 cells for the early period and M2 macrophages and T_{reg} cells for the middle period) as well as proliferation and restraint of the myogenic differentiation program.[71] Between 72 and 336 hrs after injury, another set of genes associated with muscle structural elements, cellular migration and niche remodeling peaked in expression and displayed changes in chromatin architecture. Detection of these enrichments yields a portrait whereby during this time, proliferating satellite cells migrate into the damaged area and position themselves via cues from the ECM and then proceed to differentiate and fuse into myotubes with each other (or with existing myofibers). During the latter half of the late period (504-672 hrs), the chromatin state did not vary considerably and closely resembled the chromatin state of the uninjured tissue.

Skeletal muscle regeneration is tightly regulated through multiple pathways and an excellent example is through the IGF pathway, which serially activates PI3K, Akt and mammalian target of rapamycin (mTOR). As was noted above, IGF and PI3k signaling in regenerating muscle was tightly regulated across multiple levels and temporally, especially during the transition between

myoblast proliferation and differentiation. In the early period, IGF1 was upregulated and quickly followed by increases in expression and chromatin remodeling of myogenic negative regulators (MG53, HDACs, XBP1, USF1) that alter how the myogenic regulatory factors bind to lineage-specific genes. Recent *in-vitro* results tracking MyoD binding to DNA through myogenic differentiation showed this TF exhibits dynamic binding patterns at enhancers that are modulated by SMADs and TGF-β signaling[41], which was also upregulated during this time period. Starting at 72hrs, IGF1 and negative regulators began to decrease in expression while IGF2, IGFBP5, Class IA PI3K genes, Akt1, H19, linc-MD1, miR-1a, miR-206, miR-31, miR-126b, miR-483 and miR-21 increased in expression, which is consistent with previous observations.[7,10,46,47,48]. This switch is consistent with the view that miRs-206, 29 and 1a bind to HDACs that reduce the binding affinity of MyoD and Mef2, permitting these TFs to initiate myogenic regeneration. Given IGF-activated Akt1 modulates p300 to associate with MyoD[72], it is intriguing to hypothesize how other parallel regulatory schemas of this signaling pathway alter fundamental changes in TF-binding and chromatin states that drive muscle regeneration after severe trauma. Collectively, these experiments imply muscle repair and regeneration uses different sets of transcriptional programs, noncoding RNAs, combinations of TFs as well as chromatin remodeling factors to precisely execute stage-specific gene-expression programs.

The different types of interactions viewed here offer insights into potential therapies of regenerating muscle through discrete stages and suggest chromatin restructuring therapeutic modalities may be effective to coordinate or augment cellular state transitions through integration of unique sets of chromatin remodelers and TFs. In summary, this study represents one of the first studies to monitor the in-vivo dynamics of muscle regeneration at genome-wide levels. The simultaneous measurement of epigenetic marks and RNA sequencing including coding and noncoding RNA recapitulated some of the newest players in muscle biology and allowed unbiased views of the actions of various TFs on their targets. We envision that a wider combinatorial interrogation of such a dataset can represent a valuable resource to extend the networks acting in such a complex micro-environment like the cis-regulatory modules engaged by TFs, miRNAs and lincRNAs.

Materials and Methods

Animals and Traumatic Injury Model.

Male C57BL/6J mice (10 weeks of age, 24-27 grams) were obtained from The Jackson Laboratory (Bar Harbor, ME). Mice were housed one per cage (shoebox cage, 7" x 11" x 5"h) in the USARIEM animal facility at a constant temperature (24 ± 1ºC), 50 % relative humidity, with a 12h/ 12h (0600-1800 h) light/dark cycle. Standard laboratory rodent chow and water were provided *ad libitum*. Cages were supplied with Alpha-dri and cob blend bedding for nesting and enrichment and plastic houses for warmth and comfort. Food intake and body mass were recorded daily. Mice were cared for in accordance with the Guide for the Care and Use of Laboratory Animals in a facility accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC).

Prior to administration of the freeze injury, mice were anesthetized with a combination of fentanyl (0.33 mg/kg), droperidol (16.7mg/kg), and diazepam (5 mg/kg). The TA muscle was exposed via a 1 cm long incision in the aseptically prepared skin overlying the TA muscle. Freeze injury was performed in the left, hind limb while the non-injured contralateral leg served as the control. Freeze injury was induced by applying a 6 mm diameter steel probe (cooled to the

temperature of dry ice, -70C) to the belly of the TA muscle (directly below incision site) for 10 seconds. Following injury, the skin incision was closed using 6-0 plain gut absorbable suture (Ethicon, Piscataway, NJ). The analgesic, Buprenorphine (0.1 mg/kg SQ) was administered using a 25-27 gauge needle prior to recovery from anesthesia.

Five mice were euthanized at each time-point post-injury (3, 10, 24, 48, 72, 168, 336, 504, 672 h) via CO_2 inhalation (2 liters/min), thoracotomy and exsanguination. TA muscles were removed from the injured and contralateral limb; weighed, and approximately half of the tissue was processed further (area delineated by the injury). 80-85% of the extracted tissue was crosslinked in 1% formaldehyde for 15 min at room temperature, with constant agitation. The reaction was then quenched with 125mM Glycine for 5 min at room temperature with constant agitation. The fixed tissue was then rinsed 3x with cold PBS (4ºC), spun down, snap frozen in liquid nitrogen and stored for future use.

Satellite Cell Isolation and Fluorescence-Activated Cell Sorting.

Muscle satellite cells were isolated from cryo-injured quadriceps and tibialis anterior muscles, as described previously[73,74] at 3, 24, 48, 72, and 168 hours post-injury. Briefly, intact muscles were harvested from C57BL/6J mice and digested in 0.2% collagenase type II in DMEM at 37°C for 1.5 hours. Upon digestion, muscles were minced to dissociate muscle fibers. Muscle fibers were further digested in F10 media containing 0.0125% collagenase type II and 0.05% dispase at 37°C for 30 min. Isolated cells were incubated in Hank's Buffered Salt Solution containing 2% donor bovine serum and antibody cocktail (1:100) consisting of the following: allophycocyanin (APC)-conjugated anti-mouse Sca-1 (13.3% of cocktail volume, Biolegend), CD45 (13.3%, Biolegend), Mac-1 (13.3%, Biolegend), and Ter-119 (13.3%, Biolegend); phycoerythrin (PE)-conjugated anti-mouse beta-1 integrin (20.0%, Biolegend); biotinylated anti-mouse CD184 (26.8%, BD Pharmigen); and PE-Cy7-conjugated streptavidin (1:100, eBioscience), for 20 min on ice. Muscle satellite cells were identified as propidium iodide$^-$, Sca-1$^-$, CD45$^-$, Mac-1$^-$, Ter-119$^-$, β1-integrin$^+$, CD184$^+$ and sorted via fluorescence-activated cell sorting (BD FACS Aria IIIu Cell Sorter) directly into Trizol (Thermo Fisher). After sorting, each solution was snap-frozen on dry-ice and stored at -80°C until processing.

Chromatin Isolation and Sequencing Library Preparation.

Each frozen, cross-linked tissue was thawed for at least 30 min and homogenized (Tissue Ruptor, Qiagen) for 30 seconds. The slurry was then resuspended in lysis buffer (20mM Tris-HCl, 85mM KCl, 0.5% NP-40) with protease inhibitor tablets (Roche) for at least 10 minutes at 4ºC. The suspension was then centrifuged (2500 x g for 5 minutes), supernatant removed and pellet was re-suspended in nuclei lysis buffer (10mM Tris-HCl, 2% NP-40, 0.1% SDS, 0.5% 0.5% sodium deoxycholate) for at least 10 minutes at 4ºC. The chromatin was then sheared for 10 minutes using a Branson sonifier (40% amplitude, 0.7 seconds on and 1.3 seconds off) into fragments of size range of 150-700 base-pairs and let sit for 10 minutes at 4ºC. After incubation, the samples were spun down at 15000 x g for 10 minutes at 4ºC and the supernatant was collected. The supernatant was split into three separate micro-centrifuge tubes and diluted with ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM TRis-HCL, 167mM NaCl) with protease inhibitor tablets to 1.8mL. Each vial of sheared chromatin was then incubated overnight at 4ºC with a validated antibody (Modified Histone Peptide Array – Active Motif) for H3K4me3 – Millipore Catalog # 07-743, H3K27ac – Active Motif Catalog # 39133, or H3K4me1 – Abcam Catalog # ab8895 using constant agitation. The antibody-chromatin complexes were

immunoprecipitated using a mixture of Protein A & Protein G Dynabeads (Life Technologies) for 2 hrs at 4ºC with constant agitation. The immunoprecipitated chromatin was then re-suspended in low salt wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, 150mM NaCl), incubated on ice for 5 minutes followed by re-suspension in high salt wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 500mM NaCl) and incubation on ice for 5 minutes. The samples were then re-suspended in LiCl wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, 150mM NaCl) and incubated on ice for 5 minutes. After incubation in LiCl, each sample was re-suspended in TE buffer (10mM Tris-HCl, 1mM EDTA) at room temperature and incubated for 5 minutes. Each re-suspension in low-salt wash buffer, high-salt wash buffer, LiCl wash buffer or TE buffer was repeated and supernatant aspirated. Lastly, each sample was re-suspended in ChIP elution buffer (1% SDS, 0.1M $NaHCO_3$) and incubated at 65ºC for 15 minutes. The collected supernatant was then incubated overnight in a reverse cross-linking salt mixture (250mM Tris-HCl, 62.5 mM EDTA, 1.25m NaCl, 5 mg/mL proteinase K) at 65ºC. DNA was purified with solid-phase reversible immobilization (SPRI) magnetic beads (Agencourt), washed with 80% ethanol, eluted and quantified (Qubit, Invitrogen). Approximately 10ng of isolated DNA from each ChIP was end-repaired (End-It DNA End-Repair Kit, Epicenter), extended and A-tailed (New England Biolabs), and ligated to sequencing adaptor oligos (Illumina). The adaptor-modified library was then amplified by PFU Ultra II Hotstart Master Mix (Agilent) and size-selected to a range of 300-600 base-pairs prior to sequencing. Libraries were pooled and sequenced using an Illumina Genome Analyzer IIx using 44 base-pair single end reads to achieve approximately 20 million aligned reads per sample.

mRNA and Small-RNA Sequencing Library Preparation and Quantitative PCR.

A portion of the extracted tissue was snap frozen in Trizol (Thermo Fisher) and thawed, homogenized for 30 seconds (Tissue Ruptor, Qiagen) and total RNA was isolated from the tissue after homogenization using the miRNeasy Mini Kit (Qiagen) as per the manufacturer's instructions. Total RNA from sorted satellite cells were extracted using the miRNeasy Micro Kit (Qiagen) as per the manufacturer's instructions and RNA concentration and integrity for both types of samples (tissues and sorted cells) were measured with a Nanodrop spectrophotometer (Nanodrop 2000c) and Bioanalyzer (Agilent 2100). If a sample did not pass quality metrics for further processing (RIN>7), the samples were omitted from further processing. For the RNA isolated from the tissues, 1 μg of isolated total RNA was used to produce strand-specific cDNA libraries using the Truseq (Illumina) protocol, as per the manufacturer's instructions and previously described.[16] For the RNA isolated from the satellite cells, 10ng of isolated total RNA was used to produce cDNA libraries using the SmartSeq4 protocol (Clontech), as per the manufacturer's instructions. Individual libraries from tissues were pooled and sequenced using twelve lanes of 76-bp paired-end reads on an Illumina Genome Analyzer IIx to an average depth of 50M reads per library. Individual libraries for the sorted satellite cells were pooled and sequenced using 76-bp paired-end reads on an Illumina NextSeq to an average depth of 50M reads per library.

To produce small-RNA sequencing libraries, 500 ng of total RNA from the tissues and 100 ng of total RNA from the sorted cells was used and libraries were prepared according to the specifications of the Truseq small-RNA kit (Illumina). Each library was then pooled and sequenced on a MiSeq single ended 35bp run (Illumina) to >1M reads per library. Several miRNAs observed from the miRNA-Seq dataset were validated using miRNA-qPCR and the TaqMan MicroRNA Reverse Transcription Kit (Catalog # 4366596). Primers for miR206

(Taqman-Thermo Fisher Catalog # 4426961) and miR133a (Taqman-Thermo Fisher Catalog # 4426961) were used to perform quantitative PCR and the comparative dCt method.

Long Non-Coding RNA (lncRNA) Data Processing.

The bedtools software suite[75] was used to detect transcripts aligned to long non-coding RNA (lncRNA) loci associated with myocyte development, in RNA-seq files. BAM files were converted to BED format using the bamtobed utility. BED files were sorted by chromosome then transcript start position, relative to the alignment genome (mm9) using the UNIX sort utility (sort –k1,1 –k2,2n <filename>). The bedtools intersect utility was used to search the sorted BED files for mRNA transcripts which overlap with lncRNA loci. Only transcripts with 100% overlap with the reference loci were included in downstream analysis and lncRNA expression was normalized to the lncRNA length and visualized using the R heatmap library. Long non-coding RNA differential expression analysis was performed within the R statistical environment, using bioinformatics packages "DESeq2", "Annotationdbi", "BiocParallel", "GenomicAlignments", "pheatmap", and "GenomicFeatures" from Bioconductor. A Gene x Sample count matrix was generated by the summarizeOverlap function (GenomicAlignments package) from paired-end read bam file transcripts mapped to the UCSC mm9 Refgene annotation file. The annotated count matrix gene list was then filtered to contain only transcripts with the Refseq non-coding "NR_" prefix in the identifier and length greater than 200 bp. Out of 3029 total LncRNAs in the UCSC Refgene reference list, 2444 were identified in the experimental bam files.

Differential expression of the identified LncRNAs across and within 8 time points relative to a baseline (672 hours after injury) was estimated using a time series comparison general linear model (design = ~ Timepoint(in hrs)+ Condition + Timepoint:Condition) computed by DESeq (DESeq2 package). LncRNAs with a $\log_2$ fold change in expression of less than 2 were removed from the analysis. Significance of differential expression was assessed by Benjamini-Hochberg FDR < 0.1. Filtering of identified LncRNAs by $\log_2$ fold change > 2 and an FDR < 0.1 produced a list of 124 LncRNAs.

MicroRNA (miRNA) Data Processing.

A custom program was used to match the FASTQ sequences to the microRNA database (miRBase, www.mirbase.org) for the mouse species (mm9). The program counts exact matches and matches with 17+ base identities with two or fewer mismatches as reads. The program then tallies the reads for each miRNA detected and DESeq was used to normalize the miRNA counts. From the 8 miRNA samples (3h 28d Right, and 3h, 10h, 72h, 7d, 14d, 28d), 35 miRNAs were selected as being differentially expressed. Next, targets of the 35 miRNAs were extracted from microRNA database (cutoff alignment 140, energy -7, mirSVR -0.1) if they were in the differentially expressed RNA-Seq list (at least one sample with average FPKM >5 and fold change >3 over control at any time point). A pairwise correlation for miRNAs and mRNAs was conducted in R using the Pearson correlation coefficient with FDR correction (Benjamini-Hochberg) for p-values > 0.05. Results were clustered using single–linkage hierarchical clustering (Manhattan metric) and heat map generation in Multi–experiment Viewer.

ChIP-Seq Data Processing.

Individual replicates were aligned to the mm9 mouse genome using Bowtie2[76]. Replicates for a single mark at a single timepoint were merged with the samtools[77] cat command. MACS2[78] was used to call peaks for individual replicates (versus the merged controls for a given timepoint) as well as for the merged datasets (versus the merged controls for a given timepoint). The fold change parameter in MACS2 was set to the [2,500] range, and the –broadpeaks option was utilized to call peaks for the K4me1 datasets. The –to-large parameter was used to scale the smaller dataset to the larger dataset and the extension size was set to 200 for the K4me1 datasets.

The standard workflow for the IDR framework[79] was used to identify irreproducible peaks from individual replicates: any peaks that were not present in the majority of replicates were excluded from the set of peaks identified by MACS2 for the merged datasets. The peaks that remained after the IDR filtering were analyzed with the GREAT toolkit[80] for the mm9 genome to identify genes and pathways associated with the enriched sites. The gene regulatory domain definition was used to set the association rules. For the K4me3 marks, the GREAT association rules were set to analyze proximal regions 1.5 kbp upstream, 1.5 kbp downstream, and distal 1.5 kbp. For the remaining two marks, the association rules were set to analyze proximal regions 5.0 kbp upstream, downstream, and distally. The MSigDB pathways identified by GREAT with FDR <0.05, using the Bonferroni correction, were selected. Pathways with significant FDR values at 8 timepoints for K27ac, 7 timepoints for K4me3, and 5 timepoints for K4me1 were identified, and significance levels were plotted over time.

In a parallel analysis, the post-IDR peaks from the merged replicates were compared to the genes.gtf file for mm9. Peaks were mapped to promoter regions (within 1.5 kbp of the TSS), the first exon, subsequent exons, and introns, as determined from the exons.gtf file. Peaks were additionally mapped to the near-gene region, defined as within 5 kbp upstream of the TSS or 5 kbp downstream from the final exon. Remaining peaks were assigned to the intergenic region.

The HOMER software suite[81] was used to discover known and *de novo* motifs for the K27ac, K4me3, and K4me1 datasets. The findMotifsGenome.pl tool within the HOMER toolkit was used with the mm9 reference genome, and a threshold of 1e-50 was used to determine result significance for de novo motif discovery.

The final set of peaks from the merged datasets were compared against the MyoD ChIP-Seq datasets from Cao et al[14] (GSM857390, GSM857391) as well as a MyoG ChIP-Seq dataset from ENCODE[15] (ENCFF001XVD). The bedtools intersection tool was used to identify overlapping peaks across these datasets and the datasets generated as part of this study. Upregulated promoter regions were confirmed by analyzing the overlap between the Myod/Myog datasets and the K4me3 dataset. The K4me3 dataset was then compared against the K27ac and K4me1 datasets to identify and remove common peaks, leaving only K27ac and K4me1 peaks that map to enhancer regions. These remaining peaks were subsequently also overlapped with the MyoD/MyoG datasets to confirm upregulated enhancer regions.

The KEGG and GO[82] database was queried to identify pathways that were over-expressed in the injured samples. Additionally, Gene Set Enrichment Analysis[83] using the MsigDB v. 4.0 was performed to identify differentially enriched gene sets.

Data Visualization

The IGV browser (v.2.3.36)[84] was used to visualize pileup data and to generate the tracks in Figures 1-6. All chromatin maps were converted and visualized as bigWig files normalized to 10M reads.

Acknowledgments

The authors thank Chet Beal for assistance with artwork, Sara Chauvin, Patrick Boyle and Fontina Kelley and the Broad Institute Genomics Platform for sequencing and technical assistance, Tara Boettcher for assistance with ChIP and sequencing library preparation, Mary Abdalla, Christina Zook and Alyssa Geddis for technical assistance with PCR, and Darrell O. Ricke and Michael J. Ziller for insightful discussions. This material is based upon work supported under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Opinions, interpretations, recommendations and conclusions are those of the authors and are not necessarily endorsed by the United States Government. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Air Force. C.C. was supported by an appointment to the postgraduate research participation program at the US Army Research Institute of Environmental Medicine, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the US Army Medical Research and Materiel Command. The views, opinions, and/or findings of this report are those of the authors and should not be construed as an official US Department of the Army position, policy, or decision unless so designated by other official documentation.

Author contributions

C.A.A., M.L.U. and A.M. conceived the study, C.A.A., M.L.U., R.P., D.C., C.T.C., C.A.G., M.A.K., W.M.H., E.J.S., Y.C.J., R.W.M. performed experiments, C.A.A., A.S., A.W., J.L. performed computational analysis of the data, C.A.A. wrote the paper with assistance from the other authors.

Competing interests

The authors declare no competing financial interests.

Figures and Tables

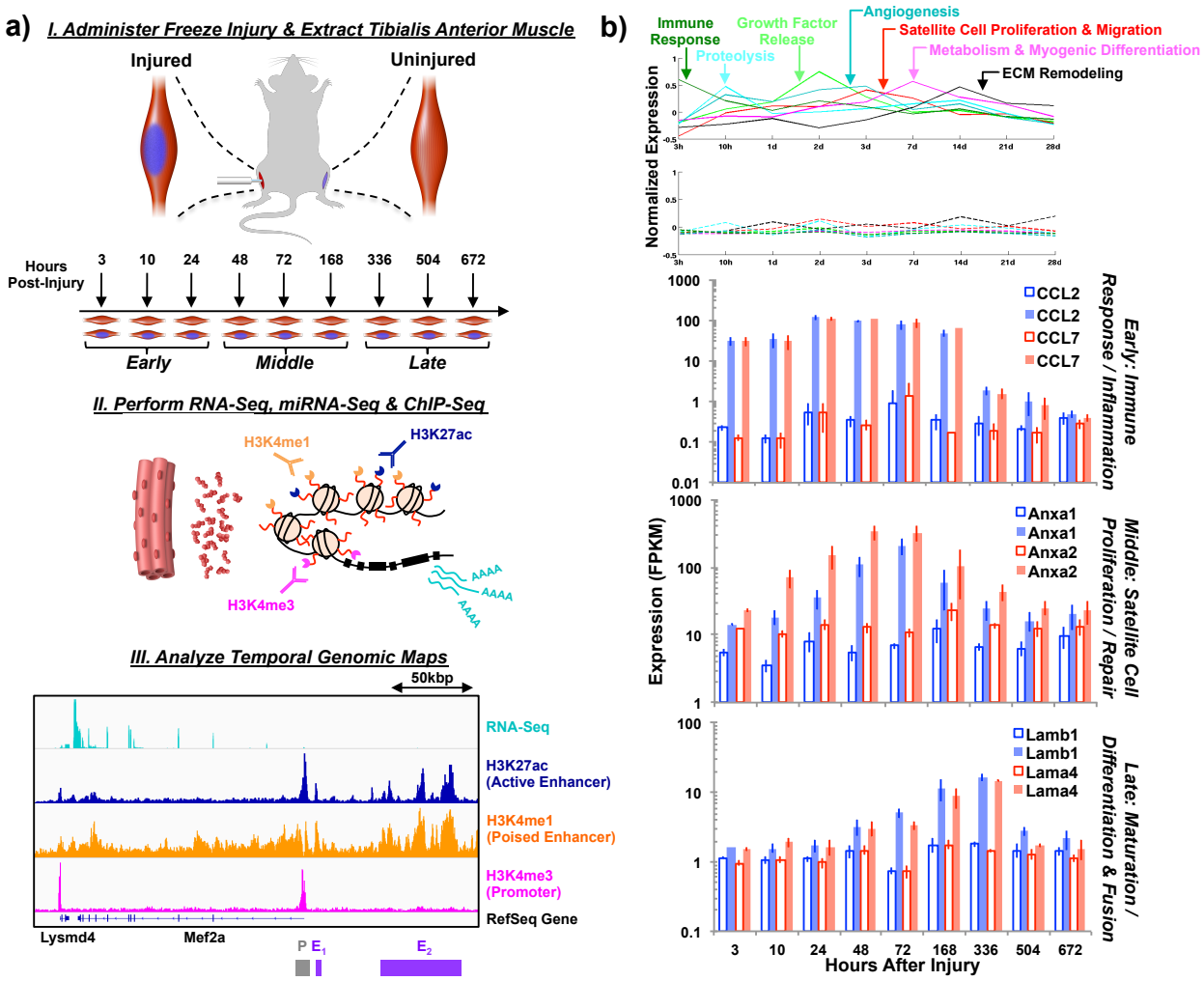

Figure 1. Experimental overview for profiling molecular mechanisms governing *in-vivo* tibialis anterior (TA) muscle regeneration after severe trauma. a) Schematic diagram of injury model and process flow for chromatin and transcript extraction. In I, the TA muscle of a C57BL/6J mouse is administered a freeze injury and the contralateral TA muscle serves as the control. The injured and uninjured tissues are extracted at different times after the injury was administered (3hrs-672 hrs). During II, the tissue is digested and polyadenylated RNA and small RNA as well as different chromatin modifications are extracted and enriched. In III, the enriched nucleic acids are then sequenced, aligned to the mm9 genome and analyzed. A representative example of the Mef2a gene at 3 hours post-injury is shown where the promoter (labeled P in gray) and enhancer regions (labeled E1 and E2 in purple) are depicted. b) (Top) Line plots of hierarchically clustered RNA-Seq data through time revealed clusters up- and downregulated at different time periods that were associated with different stages of muscle repair and regeneration. The top panel is for the injured datasets and the bottom panel is for the contralateral uninjured datasets. (Bottom) Bar graphs of gene expression values of six different genes corresponding to different stages of the muscle regeneration process through time from left to right.

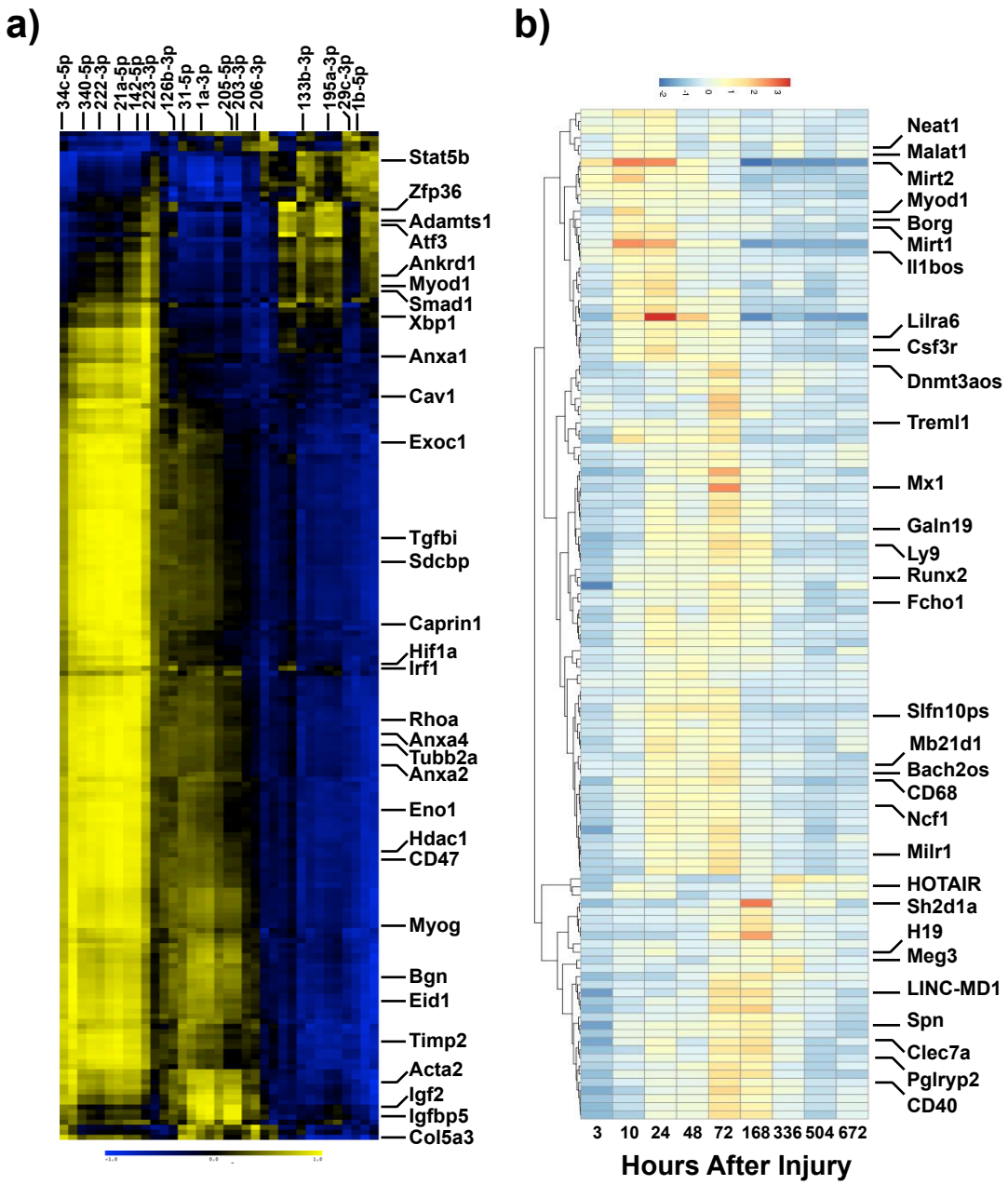

Figure 2. Dynamics of non-coding RNAs after severe muscle trauma. a) Cross-correlation analysis of dynamic miRNAs and their mRNA targets. The heatmap is plotted as the Pearson correlation coefficient of the expression of 35 dynamic miRNAs (x-axis) against the expression of their 200 predicted mRNA targets (y-axis). A positive correlation coefficient is labeled in yellow and a negative correlation coefficient is labeled in blue. b) Expression heatmap of 124 long noncoding RNAs through time from left to right. Heatmap is depicted as fold change of injured vs uninjured where blue is a negative fold change and red is a positive fold change.

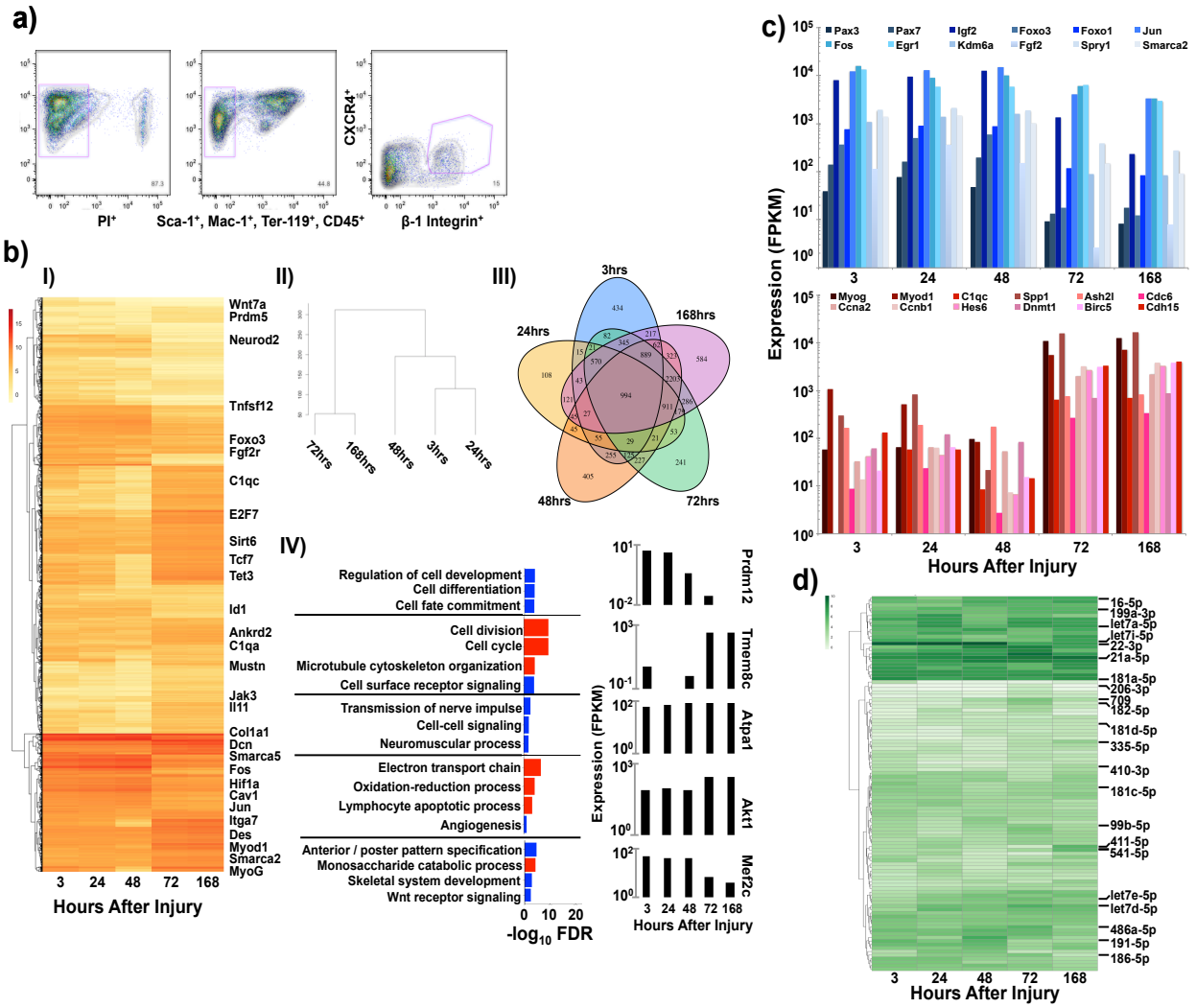

Figure 3. Temporal coding and non-coding transcriptional signatures of sorted satellite cells post-trauma highlights regenerative transitions. a) Representative isolation plot of Fluorescence-Activated Cell Sorting (FACS) of satellite cells for Sca-1$^-$, CD45$^-$, Mac-1$^-$, Ter-119$^-$, β1-integrin$^+$, Cxcr4$^+$. b) I) Expression heatmap of 994 representative differentially expressed genes through time where yellow is low expression and red is high expression and several representative genes are listed. II) Dendrogram showing global hierarchical clustering of RNA-Seq datasets separated by their Jensen-Shannon distance. III) Venn diagram of unique and overlapping differentially expressed genes for the time points sampled. IV) Bar graphs of False Discovery Rates (FDR) for overrepresented Gene Ontology (GO) pathways derived from hierarchical clustering of the genes into 5 clusters. Blue bars are upregulated GO terms and red bars are downregulated GO terms. Bar plots of RNA expression profiles of genes from each cluster are plotted to the right of each cluster through time from left to right. c) Bar graphs of individual gene expression values through time from left to right of a cluster of genes associated with quiescent satellite cells that displayed a relatively high level of expression from 3-48hrs post-injury and decreased from 72-168 hrs (blue - top panel). The bottom panel shows a set of genes from a different cluster that was lowly expressed from 3-48hrs and upregulated from 72-168 hrs and was linked to satellite cell proliferation and differentiation (MyoG, Hes6, Ash2l, Cdc6). d) Expression heatmap of 107 differentially expressed miRNAs observed where white is low expression and green is high expression and multiple miRNAs are listed.

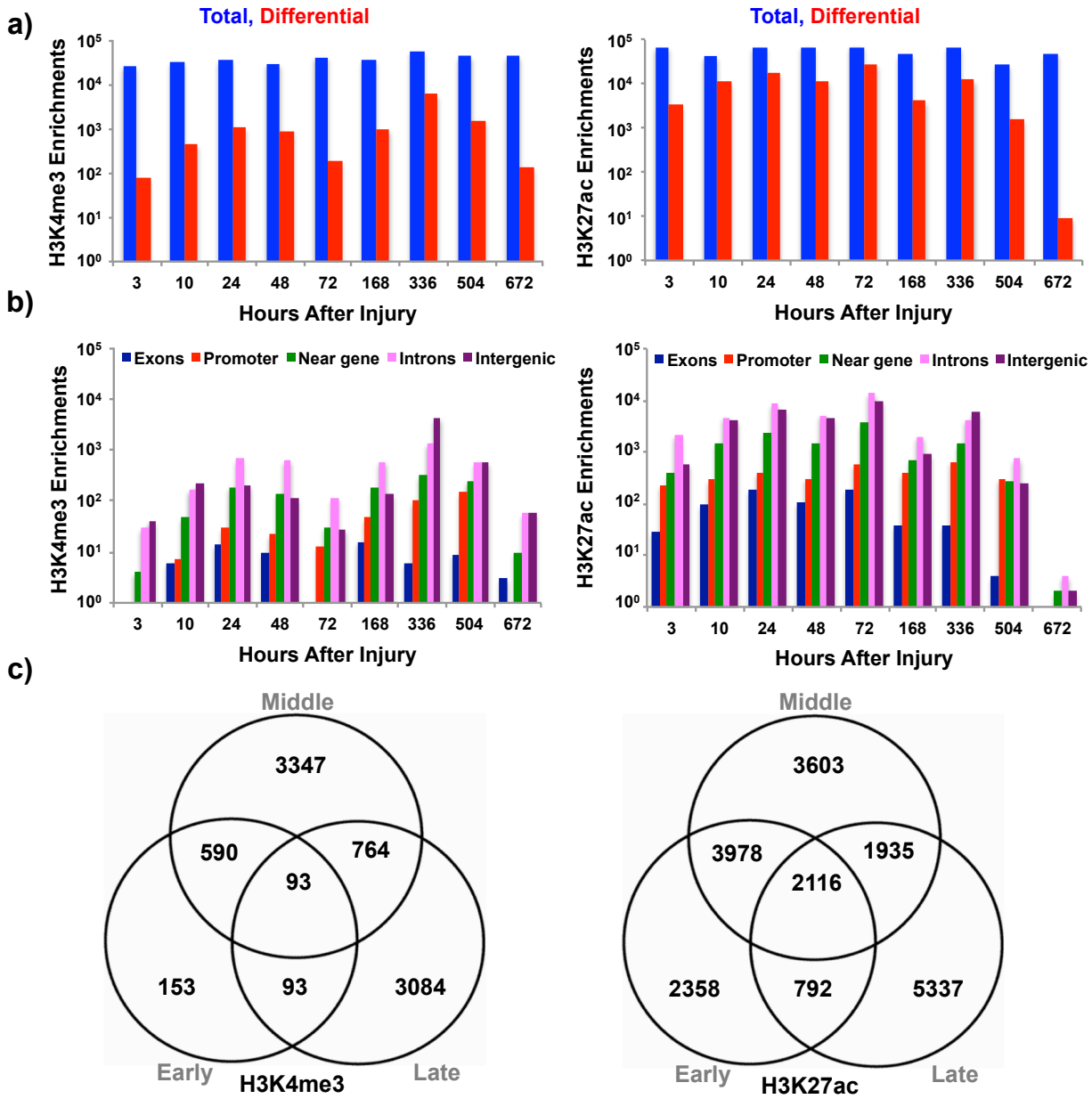

Figure 4. Distribution of enriched chromatin sites across the genome during *in-vivo* muscle regeneration. a) Total and differential number of enriched sites for H3K4me3 and H3K27ac. b) Distribution of peaks across various genomic elements (promoter, intergenic, intron, near-gene: ± 2kbp from transcriptional start site, exon). c) Shared and unique numbers of enriched chromatin sites corresponding to different genomic elements in b).

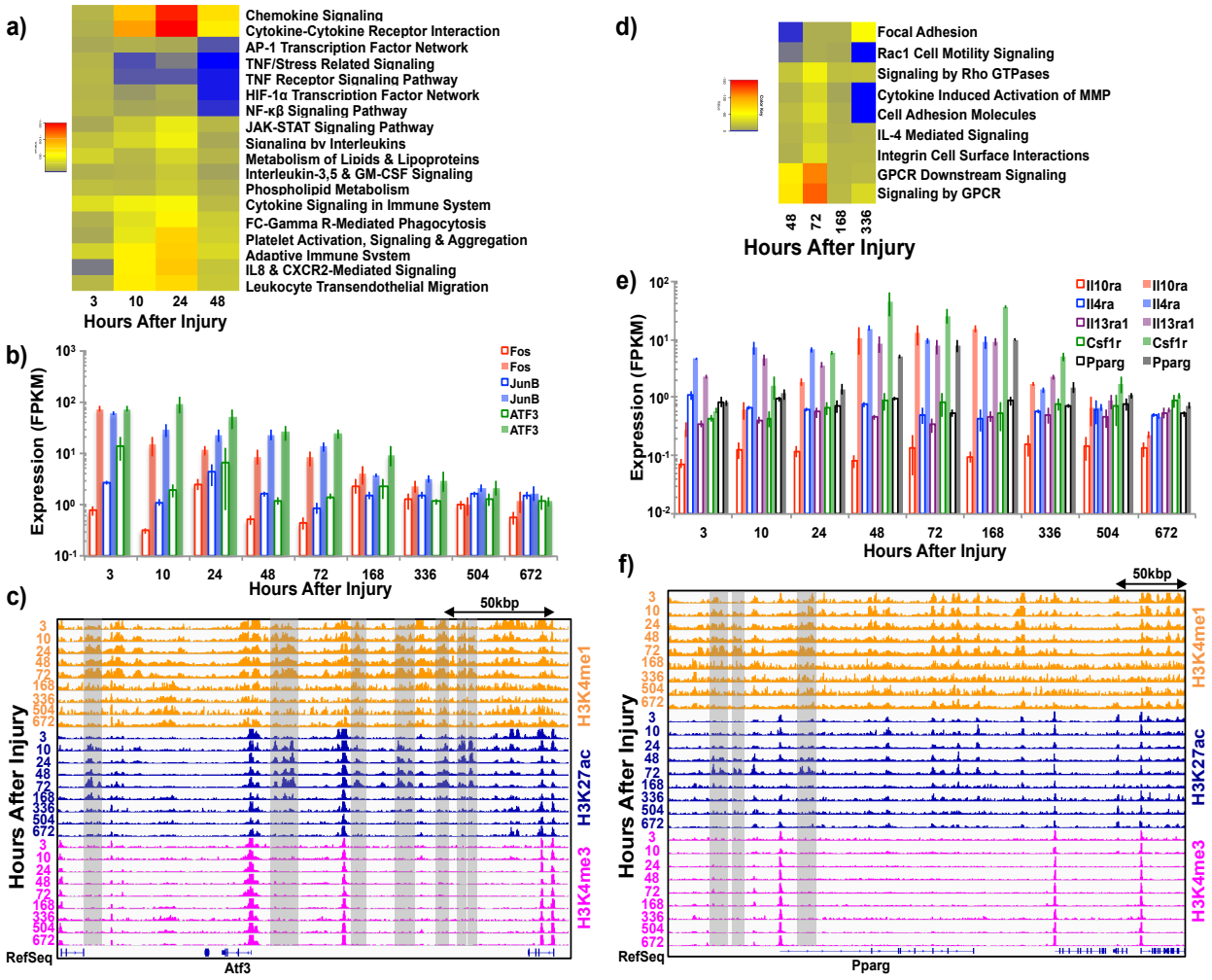

Figure 5. Chromatin landscapes are immediately modified after severe muscle trauma and reflect recruitment of different types of immune cells. a) Heatmap of p-values for overrepresented pathways derived from enriched H3K27ac peaks in the early time period (3–48 hrs). b) Bar graphs of individual gene expression values through time from left to right of immune-cell related transcripts upregulated in the early period (injured samples are colored and uninjured samples are uncolored). c) Normalized ChIP-Seq tracks of H3K4me1, H3K27ac, and H3K4me3 profiles around the Atf3 gene, a component of the AP-1 complex. Enriched enhancer regions are highlighted in gray. d) Heatmap of overrepresented pathways derived from enriched H3K27ac peaks in the middle time period (48–336 hrs). e) Bar graphs of individual gene expression values through time from left to right of anti-inflammatory related transcripts. Interleukin-4 (IL-4ra), interleukin-10 (IL-10ra) and interleukin-13 (IL-13ra1) induce macrophage M2 polarization and modulate MyoD expression. Colony stimulating factor 1 receptor (Csf1r) is a cytokine that mediates macrophage differentiation and peroxisome proliferator-activated receptor gamma (Ppar-γ) is a transcription factor that modulates immune cell differentiation. f) Normalized ChIP-Seq tracks of H3K4me1, H3K27ac, and H3K4me3 profiles around the Ppar-γ gene. Enriched enhancer regions are highlighted in gray.

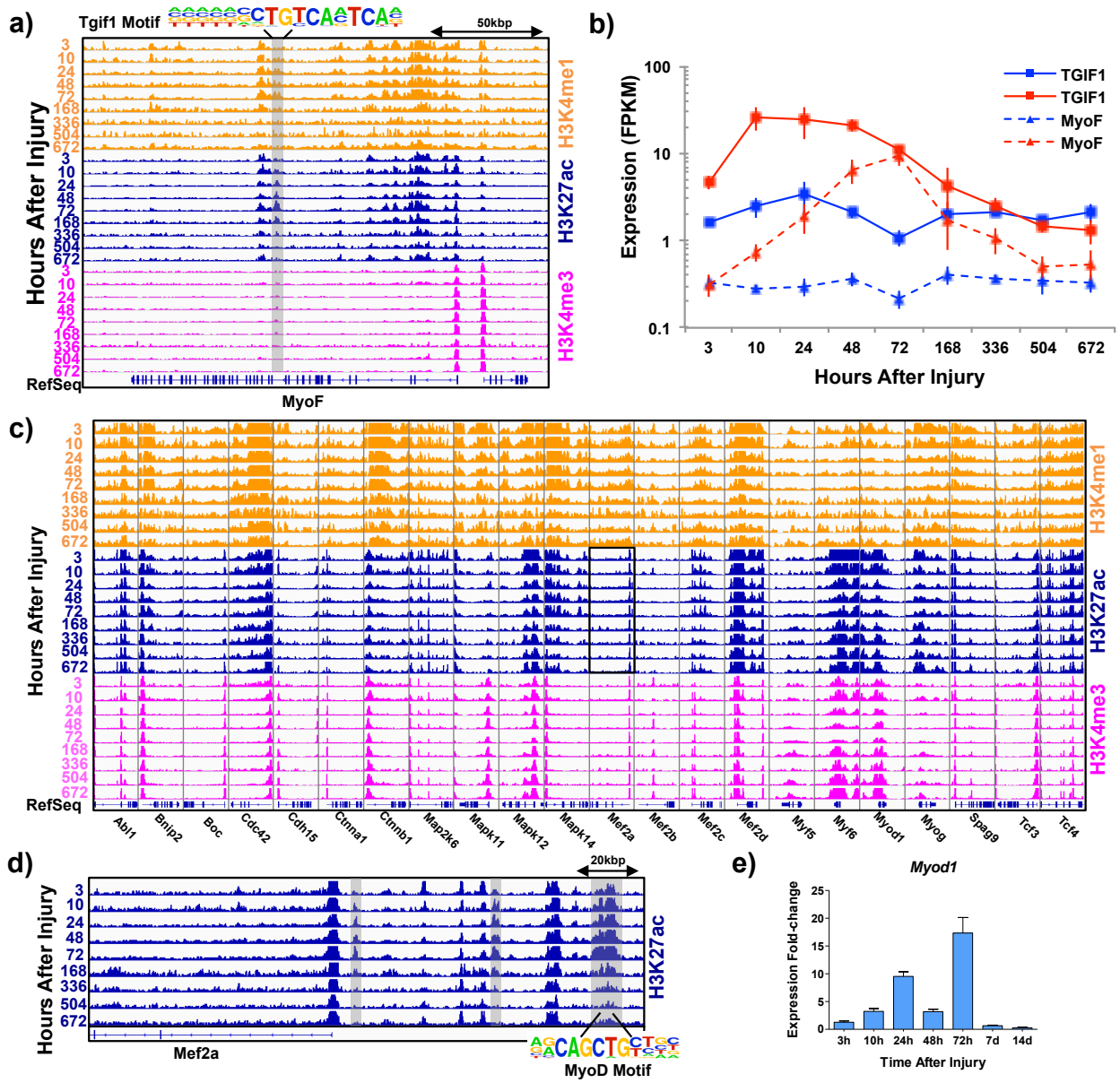

Figure 6. Chromatin state transitions associated with activation of basement membrane repair & myogenic regeneration. a) Normalized ChIP-Seq tracks of H3K4me1, H3K27ac, and H3K4me3 profiles around the myoferlin (MyoF) gene, a plasma membrane protein. Enriched enhancer regions are highlighted in gray and corresponding enriched TF motif is labeled. b) Line graphs of individual gene expression values of MyoF and associated transcription factor (TGIF1) through time from left to right (injured samples are colored red and uninjured samples are colored blue). c) Normalized ChIP-Seq tracks of H3K4me1, H3K27ac, and H3K4me3 profiles for a subset of myogenesis genes. The black box outlines the H3K27ac region which is expanded in d) and illustrates the Mef2a locus. d) Normalized H3K27ac ChIP-Seq track highlighting enriched enhancers around the Mef2a gene, a myogenic regulatory transcription factor. The enriched enhancer regions are highlighted in gray and corresponding enriched TF motif (MyoD) is labeled beneath the track. e) Bar graph of expression for MyoD for time points where a fold change was observed (504 hrs and 672 hrs did not yield changes in expression).

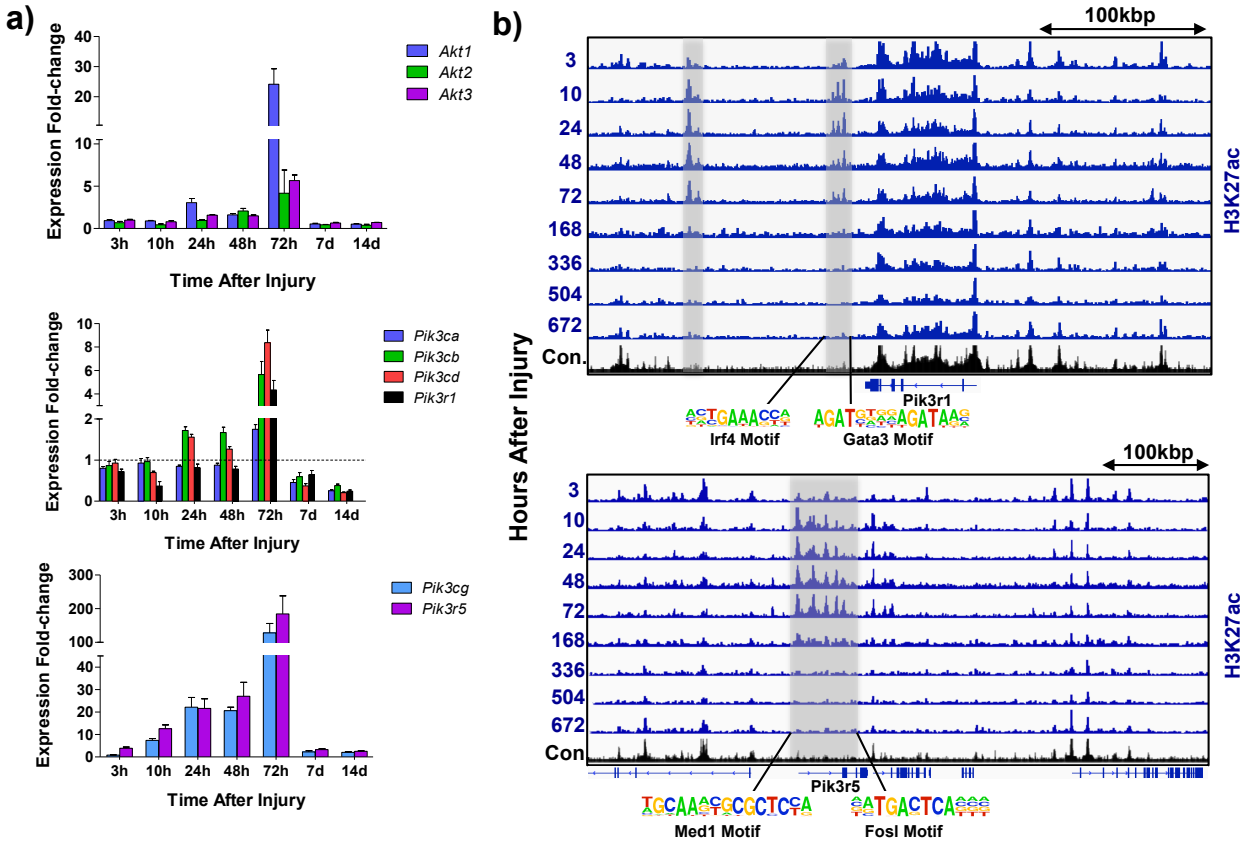

Figure 7. PI3K/Akt activation promotes transition from proliferation towards myogenic differentiation. a) Bar graphs of individual gene expression values for various components of the PI3K and AKT pathways through time from left to right. b) Normalized ChIP-Seq tracks of H3K27ac profiles showing differential enhancer activity of the PI3Kr1, PI3Kr5 and PI3Kr6 loci. Enriched enhancer regions are highlighted in gray and corresponding enriched TF motifs are labeled underneath.

References

1. Scharner, J. and Zammit, P.S. The muscle satellite cell at 50: the formative years. *Skeletal Muscle* **1**, 28 (2011).
2. Bentzinger, C.F., Wang, Y.X., Dumont, N.A., and Rudnicki, M.A. Cellular dynamics in the muscle satellite cell niche. *EMBO Rep.* **14**, 1062-1072 (2013).
3. Kuang, S., Gillespie, M.A., Rudnicki, M.A. Niche regulation of muscle satellite cell self-renewal and differentiation. *Cell Stem Cell* **2**, 22-31 (2008).
4. Aurora, A.B. and Olson, E.N. Immune modulation of stem cells and regeneration. *Cell Stem Cell* **15**, 14-25 (2014).
5. Burzyn, D., Kuswanto, W., Kolodin, D., Shadrach, J.L., Cerletti, M., Jang, Y., Sefik, E., Tan, T.G., Wagers, A.J., Benoist, C. and Mathis, D. A special population of regulatory T cells potentiates muscle repair. *Cell* **155**, 1282-1295 (2013).
6. Heredia, J.E., Mukundan, L., Chen, F.M., Mueller, A.A., Deo, R.C., Locksley, R.M., Rando, T.A. and Chawla, A. Type 2 innate signals stimulate fibro/adipogenic progenitors to facilitate muscle regeneration. *Cell* **153**, 376-388 (2013).
7. Braun, T. and Gautel, M. Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nature Rev. Molec. Cell Biol.* **12**, 349-361 (2011).
8. Asp, P., Blum, R., Vethantham, V., Parisi, F., Micsinai, M., Cheng, J., Bowman, C., Kluger, Y., and Dynlacht, B.D. Genome-wide remodeling of the epigenetic landscape during myogenic differentiation. *Proc. Nat'l. Acad. Sci. USA* **108**, E149-E158 (2011).
9. Brancaccio, A., Palacios, D. Chromatin signaling in muscle stem cells: interpreting the regenerative microenvironment. *Front. Aging Neurosci.* **7**, 1-17 (2015).
10. Buckingham, M. and Rigby, P.W.J. Gene regulatory networks and transcriptional mechanisms that control myogenesis. *Devel. Cell* **28**, 225-238 (2014).
11. Roadmap Epigenomics Consortium et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-330 (2015).
12. Liu, L., Cheung, T.H., Charville, G.W., Hurgo, B.M.C., Leavitt, T., Shih, J., Brunet, A. and Rando, T.A. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell Rep.* **4**, 189-204 (2013).
13. Giordani, L. and Puri, P.L. Epigenetic control of skeletal muscle regeneration: integrating genetic determinants and environmental changes. *FEBS* **280**, 4014-4025 (2013).
14. Cao, Y., Yao, Z., Sarkar, D., Lawrence, M., Sanchez, G.J., Parker, M.H., MacQuarrie, K.L., Davison, J., Morgan, M.T., Ruzzo, W.L., Gentleman, R.C., Tapscott, S.J. Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Dev. Cell* **18**, 662-674 (2010).
15. The Mouse ENCODE consortium. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**, 355-364 (2014).
16. Aguilar, C.A., Shcherbina, A., Ricke, D.O., Pop, R., Carrigan, C.T., Gifford, C.A., Urso, M.L., Kottke, M.A., and Meissner, A. In vivo monitoring of transcriptional dynamics after lower-limb muscle injury enables quantitative classification of healing. *Sci. Rep.* **5,** 13885 (2015).
17. Tidball, J. G. Inflammatory processes in muscle injury and repair. *J. Physiol. Regul. Integr. Comp. Physiol.* **288**, R345-R353 (2005).

18. Warren, G.L., Summan, M., Gao, X., Chapman, R., Hulderman, T., Simeonova, P.P. Mechanisms of skeletal muscle injury and repair revealed by gene expression studies in mouse models. *J. Physiol.* **582**, 825-841 (2007).
19. Chen, J.F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L., and Wang, D.Z. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature Genet.* **38**, 228-233 (2005).
20. Kim, H.K., Lee, Y.S., Sivaprasad, U., Malhotra, A., Dutta, A. Muscle specific microRNA miR-206 promotes muscle differentiation. *J. Cell Biol.* **174**, 677-687 (2006).
21. Zhang, Y., Yang, P., Sun, T., Li, D., Xu, X., Rui, Y., Li, C., Chong, M., Ibrahim, T., Mercatali, L., Amadori, D., Lu, X., Xie, D., Li, Q.J., Wang, X.F. miR-126 and miR-126* repress recruitment of mesenchymal stem cells and inflammatory monocytes to inhibit breast cancer metastasis. *Nature Cell Biol.* **15**, 284-294 (2013).
22. Cachiarelli, D., Incitti, T., Martone, J., Cesana, M., Cazzella, V., Santini, T., Sthandier, O., Bozzoni, I. miR-31 modulates dystrophin expression: new implications for Duchenne muscular dystrophy therapy. *EMBO Rep.* **12**, 136-141 (2011).
23. Cachiarelli, D., Martone, J., Girardi, E., Cesana, M., Incitti, T., Morlando, M., Nicoletti, C., Santini, T., Sthandier, O., Barberi, L., Auricchio, A., Musaro, A., Bozzoni, I. MicroRNAs involved in molecular circuitries relevant for the Duchenne muscular dystrophy pathogenesis are controlled by the dystrophin/nNOS pathway. *Cell Metab.* **12**, 341-351 (2010).
24. Liu, N., Williams, A.H., Maxeiner, J.M., Bezprozvannaya, S., Shelton, J.M., Richardson, J.A., Bassel-Duby, R., Olson, E.N. microRNA-206 promotes skeletal muscle regeneration and delays progression of Duchenne muscular dystrophy in mice. *J. Clin. Invest.* **122**, 2054-2065 (2012).
25. Neguembor, M.V., Jothi, M., Gabellini, D. Long noncoding RNAs, emerging players in muscle differentiation and disease. *Skeletal Muscle* **4:8**, 1-12 (2014).
26. Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A., Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **147**, 358-369 (2011).
27. Gong, C., Li, Z., Ramanujan, K., Clay, I., Zhang, Y., Lemire-Brachat, S., Glass, D.J. A long noncoding RNA, LincMyoD, regulates skeletal muscle differentiation by blocking IMP2-mediated mRNA translation. *Dev. Cell* **34**, 181-191 (2015).
28. Dey, B.K., Pfeifer, K., Dutta, A. The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes & Develop.* **28**, 491-501 (2014).
29. Zangrado, J., Zhang, L., Vausort, M., Maskali, F., Marie, P.Y., Wagner, D.R., Devaux, Y. Identification of candidate long non-coding RNAs in response to myocardial infarction. *BMC Genomics* **15:460** (2014).
30. Pallafacchina, G., Francois, S., Regnault, B., Czarny, B., Drive, V., Cumano, A., Montarras, D., Buckingham, M. An adult tissue-specific stem cell in its niche: A agene profiling analysis of in vivo quiescent and activated muscle satellite cells. *Stem Cell Res.* **4,** 77-91 (2010).
31. Farina, N.H., Hausberg, M., Betta, N.D., Pulliam, C., Srivastava, D., Cornelison, D., Olwin, B.B. A role for RNA post-transcriptional regulation in satellite cell activation. *Skeletal Muscle* **2:21** (2012).
32. Lu, W., You, R., Yuan, X., Yang, T., Samuel, E.L.G., Marcano, D.C., Sikkema, W.K.A., Tour, J.M., Rodriguez, A., Kheradmand, F., Corry, D.B. The microRNA miR-22 inhibits the

histone deacetylase HDAC4 to promote TH17 cell-dependent emphysema. *Nature Immuno.* **16,** 1185-1194 (2015).
33. Liu, Q., Fu, H., Sun, F., Zhang, H., Tie, Y., Zhu, J., Xing, R., Sun, Z., Zheng, X. miR-16 family induces cell cycle arrest by regulating multiple cell cycle genes. *Nucleic Acids Res.* **36,** 5391-5404 (2008).
34. Arnold, C.P., Tan, R., Zhou, B., Yue, S.B., Schaffert, S., Biggs, J.R., Doyonnas, R., Lo, M.C., Perry, J.M., Renault, V.M., Sacco, A., Somervaille, T., Viatour, P., Brunet, A., Cleary, M.L., Li, L., Sage, J., Zhang, D.E., Blau, H.M., Chen, C., Chen, C.Z. MicroRNA programs in normal and aberrant stem and progenitor cells. *Genome Res.* **21,** 798-810 (2011).
35. Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M. and Bernstein, B.E. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43-49 (2011).
36. Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z., Blecher-Goren, R., Bonstein, C., Amann-Zalcenstein, D., Weiner, A., Freidrich, D., Medrim, J., Ram, O., Cheng, C., Gnirke, A., Fisher, S., Friedman, N., Wong, B., Bernstein, B.E., Nusbaum, C., Hacohen, N., Regev, A. and Amit, A. A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Molec. Cell* **47**, 1-13 (2012).
37. Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
38. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A. and Jaenisch,. R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Nat'l. Acad. Sci. USA* **107**, 21931-21936 (2010).
39. Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S. and Natoli, G. Latent enhancers activated by stimulation in differentiated cells. *Cell* **152**, 157-171 (2013).
40. Toth, K.G., McKay, B.R., De Lisio, M., Little, J.P., Tarnopolsky, M.A., Parise, G. IL-6 induced Stat3 signalling is associated with the proliferation of human muscle satellite cells following acute muscle damage. *PLOS One* **6**, e17392 (2011).
41. Mullen, A.C., Orlando, D.A., Newman, J.J., Loven, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., Dekoter, R.P., Young, R.A. Master transcription factors determine cell-type-specific responses to TGF-β signaling. *Cell* **147**, 565-576 (2011).
42. Kuwahara, K., Barrientos, T., Pipes, G.C.T., Li, S., Olson, E.N. Muscle-specific signaling mechanism that links actin dynamics to serum response factor. *Mol. Cell Biol.* **25**, 3173-3181 (2005).
43. Calve, S., Odelberg, S.J., Simon, H.G. A transitional extracellular matrix instructs cell behavior during muscle regeneration. *Dev. Biol.* **344**, 259-271 (2010).
44. Mousavi, K., Zare, H., Dell'Oso, S., Grontved, L., Gutieerz-Cruz, G., Derfoul, A., Hager, G.L., Sartorelli, V. eRNAs promoted transcription by establishing chromatin accessibility at defined genomic loci. *Molec. Cell* **51**, 606-617 (2013).
45. Palacios, D., and Puri, P.L. The epigenetic network regulating muscle development and regeneration. *J. Cell Physiol.* **207**, 1-11 (2006).

46. Pelosi, L., Giacinti, C., Nardis, C., Borsellino, G., Rizzuto, E., Nicoletti, C., Wannenes, F., Battistini, L., Rosenthal, N., Molinaro, M. and Musaro, A. Local expression of IGF-1 accelerates muscle regeneration by rapidly modulating inflammatory cytokines and chemokines. *FASEB J.* **21**, 1393-1402 (2007).
47. Schiaffino, S., Mammucari, C. Regulation of skeletal muscle growth by the IGF1-Akt/PKB pathway: insights from genetic models. *Skeletal Muscle* **1**, 1-14 (2011).
48. Matheny, R.W., and Adamo, M.L. PI3K p110 alpha and p110 beta have differential effects on Akt activation and protection against oxidative stress-induced apoptosis in myoblasts. *Cell Death Differ.* **17**, 677-688 (2010).
49. Matheny, R.W., Riddle-Kottke, M.A., Leandry, L.A., Lynch, C.M., Abdalla, M.N., Geddis, A.V., Piper, D.R., and Zhao, J.J. Role of phosphoinositide 3-OH kinase p110beta in skeletal myogenesis. *Mol. Cell Biol.* **35**, 1182-1196 (2015).
50. Matheny, R.W., Lynch, C.M., and Leandry LA. Enhanced Akt phosphorylation and myogenic differentiation in PI3K p110beta-deficient myoblasts is mediated by PI3K p110alpha and mTORC2. *Growth Factors* **30**, 367-384 (2012).
51. Yi, J.S., Park, J.S., Ham, Y.M., Nguyen, N., Lee, N.R., Hong, J., Kim, B.W., Lee, H., Lee, C.S., Jeong, B.C., Song, H.K., Cho, H., Kim, Y.K., Lee, J.S., Park, K.S., Shin, H., Choi, I., Lee, S.H., Park, W.J., Park, S.Y., Choi, C.S., Lin, P., Karunasiri, M., Tan, T., Duann, P., Zhu, H., Ma, J. and Ko, Y.G. MG53-induced IRS-1 ubiquitination negatively regulates skeletal myogenesis and insulin signalling. *Nature Commun.* **4**:2354 (2013).
52. Acosta-Alvear, D., Zhou, Y., Blais, A., Tsikitis, M., Lents, N.H., Arias, C., Lennon, C.J., Kluger, Y., and Dynlacht, B.D. XBP1 controls diverse cell type- and condition-specific transcriptional regulatory networks. *Mol. Cell* **27,** 53-66 (2007).
53. Soleimani, V.D., Yin, H., Jahani-Asl, A., Ming, H., Kockx, C.E.M., van Ijcken, W.F.J., Grosveld, F., Rudnicki, M.A. Snail regulates Myod binding-site occupancy to direct enhancer switching and differentiation-specific transcription in myogenesis. *Mol. Cell* **47**, 457-468 (2012).
54. Liu, L., Cheung, T.H., Charville, G.W., Rando, T.A. Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. *Nature Protoc.* **10,** 1612-1624 (2015).
55. Rodgers, J.T., King, K.Y., Brett, J.O., Cromie, M.J., Charville, G.W., Maguire, K.K., Brunson, C., Mastey, N., Liu, L., Tsai, C.R., Goodell, M.A., Rando, T.A. mTORC1 controls the adaptive transition of quiescent stem cells from G_0 to G_{Alert}. *Nature* **510,** 393-396 (2014).
56. Eisenberg, I., Eran, A., Nishino, I., Moggio, M., Lamperti, C., Amato, A.A., Lidov, H.G., Kang, P.B., North, K.N., Mitrani-Rosenbaum, S., Flanigan, K.M., Neely, L.A., Whitney, D., Beggs, A.H., Kohane, I.S., Kunkel, L.M. Distinctive patterns of microRNA expression in primary muscular disorders. *Proc. Nat'l. Acad. Sci. USA* **104**, 17016-17021 (2007).
57. Betel, D., Koppal, A., Agius, P., Sander, C., Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* **11:R90** (2010).
58. Faralli, H., Wang, C., Nakka, K., Benyoucef, A., Sebastian, S., Zhuang, L., Chu, A., Palii, C.G., Liu, C., Camellato, B., Brand, M., Ge, K., Dilworth, F.J. UTX demethylase activity is required for satellite cell-mediated muscle regeneration. *J. Clin. Invest.* **126,** 1555-1565 (2016).
59. Almada, A.E., Wagers, A.J. Molecular circuitry of stem cell fate in skeletal muscle regeneration, ageing and disease. *Nature Rev. Molec. Cell Biol.* **17,** 267-279 (2016).
60. Blais, A., Tsikitis, M., Acosta-Alvear, D., Sharan, R., Kluger, Y., Dynlacht, B.D. An initial blueprint for myogenic differentiation. *Genes & Develop.* **19**, 553–569 (2005).

61. Blum, R., Vethanthum, V., Bowman, C., Rudnicki, M., Dynlacht, B.D. Genome-wide identification of enhancers in skeletal muscle: the role of MyoD1. *Genes & Develop.* **26**, 2763-2779 (2012).
62. Liu, N., Nelson, B.R., Bezprozvannaya, S., Shelton, J.M., Richardson, J.A., Bassel-Duby, R., Olson, E.N. Requirement of MEF2A, C, and D for skeletal muscle regeneration. *Proc. Nat'l. Acad. Sci. USA* **111**, 4109-4114 (2014).
63. Molkentin, J.D., Black, B.L., Martin, J.F., Olson, E.N. Cooperative activation of muscle gene expression by MEF2 and myogenic BHLH proteins. *Cell* **83**, 2–14 (1995).
64. Arnold, L., Henry, A., Poron, F., Baba-Amer, Y., van Rooijen, N., Plonquet, A., Gherardi, R.K., and Chazaud, B. Inflammatory monocyes recruited after skeletal muscle injury switch into anti-inflammatory macrophages to support myogenesis. *J. Exp. Med.* **204**, 1057-1069 (2007).
65. Benhaddou, A., Keime, C., Ye, T., Morlon, A., Michel, I., Jost, B., Mengus, G., Davidson, I. Transcription factor TEAD4 regulates expression of myogenin and the unfolded protein response genes during C2C12 cell differentiation. *Cell Death Diff.* **19**, 220-231 (2012).
66. Umansky, K.B., Gruenbaum-Cohen, Y., Tsoory, M., Feldmesser, E., Goldenberg, D., Brenner, O., Groner, Y. Runx1 transcription factor is required for myoblasts proliferation during muscle regeneration. *PLoS Genet.* **11(8)**, e1005457 (2015).
67. Mullen, A.C., Orlando, D.A., Newman, J.J., Loven, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., Dekoter, R., Young, R.A. Master transcription factors determine cell-type-specific responses to TGFβ signaling. *Cell* **147**, 565-576 (2011).
68. Pouponnot, C., Jayaraman, L., Massague, J. Physical and functional interaction of SMADs and p300/CBP. *J Biol. Chem.* **273**, 22865-22868 (1998).
69. Puri, P.L., Avantaggiati, M.L., Balsano, C., Sang, N., Graessmann, A., Giordano, A., Levrero, M. p300 is required for MyoD-dependent cell cycle arrest and muscle-specific gene transcription. *EMBO J* **16**, 369–383 (1997).
70. Ma, K., Chan, J.K., Zhu, G., and Wu, Z. Myocyte enhancer factor 2 acetylation by p300 enhances its DNA binding activity, transcriptional activity, and myogenic differentiation. *Mol. Cell Biol.* **23**, 2893-2906 (2005).
71. Dionyssiou, M.G., Salma, J., Bevzyuk, M., Zakharyan, L., McDermott, J.C. Kruppel-like factor 6 (KLF6) promotes cell proliferation in skeletal myoblasts in response to TGF-β/Smad3 signaling. *Skeletal Muscle* **3:7** (2013).
72. Serra, C., Palacios, D., Mozzetta, C., Forcales, S.V., MOrantte, I., Ripani, M., Jones, D.R., Du, K., Jhala, U.S., Simone, C., Puri, P.L. Functional interdependence at the chromatin level between the MKK6/p38 and IGF1/PI3K/AKT pathways during muscle differentiation. *Mol. Cell* **28**, 200-213 (2007).
73. Sherwood, R.I., Christensen, J.L., Conboy, I.M., Conboy, M.J., Rando, T.A., Weissman, I.L., Wagers, A.J. Isolation of adult mouse myogenic progenitors: functional heterogeneity of cells within and engrafting skeletal muscle. *Cell* **119**, 543-554 (2004).
74. Cerletti, M., Jurga, S., Witczak, C.A., Hirshman, M.F., Shadrach, J.L., Goodyear, L.J., Wagers, A.J. Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. *Cell* **134**, 37-47 (2008).
75. Quinlan, A.R., Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinf.* **26,** 841-842 (2010)
76. Langmead, B., Salzberg, S. Fast gapped-read alignment with Bowtie 2. *Nature Meth.* **9**, 357-359 (2012).

77. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecassis, G., Durbin, R. The sequence alignment map format and SAM tools. *Bioinf.* **25,** 2078-2079 (2009).
78. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.J., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, R.M., Brown, M., Li, W., Liu, X.S. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, 137 (2008).
79. Li, Q., Brown, J.B., Huang, H., Bickel, P.J. Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat.* **5**, 1752-1779 (2011).
80. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.C., Schaar, B.T., Lowe, C.B., Wenger, A.M., Bejerano, G. GREAT improves functional interpretation of *cis*-regulatory regions. *Nature Biotech.* **28**, 495-501 (2010).
81. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., Glass, C.K. Simple combinations of lineage determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
82. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G. Gene ontology: tool for the unification of biology *Nature Genet.* **25**, 25-29 (2000).
83. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Nat'l. Acad. Sci. USA* **102**, 15545-15550 (2005).
84. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P. Integrative Genomics Viewer. *Nature Biotech.* **29**, 24–26 (2011).